



Pergamon

Insect Biochemistry and Molecular Biology 32 (2002) 1497–1506

*Insect  
Biochemistry  
and  
Molecular  
Biology*

www.elsevier.com/locate/ibmb

## Sequence of a cDNA and expression of the gene encoding a putative epidermal chitin synthase of *Manduca sexta*<sup>☆</sup>

Yu-Cheng Zhu <sup>a,d</sup>, Charles A. Specht <sup>b</sup>, Neal T. Dittmer <sup>c</sup>, Subbaratnam Muthukrishnan <sup>c</sup>, Michael R. Kanost <sup>c</sup>, Karl J. Kramer <sup>a,\*</sup>

<sup>a</sup> Grain Marketing and Production Research Center, ARS-USDA, 1515 College Avenue, Manhattan, KS 66502, USA

<sup>b</sup> Department of Medicine, Section of Infectious Diseases, Boston University, Boston, MA 02118, USA

<sup>c</sup> Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA

<sup>d</sup> Jamie Whitten Delta States Research Center, ARS-USDA, Stoneville, MS 38776, USA

Received 23 November 2001; received in revised form 15 April 2002; accepted 15 April 2002

### Abstract

Glycosyltransferases are enzymes that synthesize oligosaccharides, polysaccharides and glycoconjugates. One type of glycosyltransferase is chitin synthase, a very important enzyme in biology, which is utilized by insects, fungi, and other invertebrates to produce chitin, a polysaccharide of  $\beta$ -1,4-linked *N*-acetylglucosamine. Chitin is an important component of the insect's exoskeletal cuticle and gut lining. To identify and characterize a chitin synthase gene of the tobacco hornworm, *Manduca sexta*, degenerate primers were designed from two highly conserved regions in fungal and nematode chitin synthase protein sequences and then used to amplify a similar region from *Manduca* cDNA. A full-length cDNA of 5152 nucleotides was assembled for the putative *Manduca* chitin synthase gene, *MsCHS1*, and sequencing of genomic DNA verified the contiguity of the sequence. The *MsCHS1* cDNA has an ORF of 4692 nucleotides that encodes a transmembrane protein of 1564 amino acid residues with a mass of approximately 179 kDa (GenBank no. AY062175). It is most similar, over its entire length of protein sequence, to putative chitin synthases from other insects and nematodes, with 68% identity to enzymes from both the blow fly, *Lucilia cuprina*, and the fruit fly, *Drosophila melanogaster*. The similarity with fungal chitin synthases is restricted to the putative catalytic domain, and the *MsCHS1* protein has, at equivalent positions, several amino acids that are essential for activity as revealed by mutagenesis of the fungal enzymes. A 5.3-kb transcript of *MsCHS1* was identified by northern blot hybridization of RNA from larval epidermis, suggesting that the enzyme functions to make chitin deposited in the cuticle. Further examination by RT-PCR showed that *MsCHS1* expression is regulated in the epidermis, with the amount of transcript increasing during phases of cuticle deposition. Published by Elsevier Science Ltd.

**Keywords:** *Manduca sexta*; Tobacco hornworm; Chitin; Chitin synthase; cDNA; Gene; Amino acid sequence; Fruit fly; Glycosyltransferase; Nucleotide sequence; RNA expression; Phylogeny; Nematode; Yeast; Blow fly; *Drosophila*; Cuticle; Epidermis; Development; Transmembrane protein

### 1. Introduction

Chitin is a linear polysaccharide of *N*-acetylglucosamine residues joined by  $\beta$ -1,4 glycosidic linkages. It is found in the cell wall of fungi, the eggshell of nematodes, as well as in the shells (exoskeleton) of crustaceans from which it is extracted for commercial use. This polymer also plays an important structural role in

the cuticle and peritrophic matrix that lines the midgut of most insects. Multiple genes for chitin synthases have been isolated from yeast and other fungi (Bulawa, 1993). For example, seven chitin synthase genes have been identified from a single species of fungi (Mellado et al., 1995); however, that number may be far less in insects. With the completion of the *Drosophila* Genome Project (Adams et al., 2000), only two deduced protein sequences in the fly have similarity to fungal chitin synthases. Likewise, two putative chitin synthase genes are identifiable in the genome of the nematode, *Caenorhabditis elegans* (The *C. elegans* Sequencing Consortium, 1998). At the protein level, only fungal chitin synthases have been isolated and well characterized,

\* Corresponding author. Tel.: +1-785-776-2711; fax: +1-785-537-5584.

E-mail address: kramer@gmprc.ksu.edu (K.J. Kramer).

<sup>☆</sup> Authors Zhu and Specht contributed equally to this study.

particularly the three found in the yeast, *Saccharomyces cerevisiae* (Cabib et al., 2001).

So that appropriate levels of chitin are maintained during insect development, the synthesis and degradation of chitin by chitin synthase (EC 2.4.1.16) and chitinase (EC 3.2.1.14), respectively, need to be precisely regulated (Kramer and Koga, 1986). Although these enzymes are potential targets for pest control agents (Palli and Retnakaran, 1999; Spindler and Spindler-Barth, 1999; Cohen, 2001), there has been little commercial success in that regard in spite of the fact that both proteins have been the focus of rather intense research efforts.

More information has been generated about the properties of insect chitinases and their genes than about chitin synthases and their genes (Cohen, 1991; Koga et al., 1999; Merz et al., 1999). Insect chitinase is primarily a soluble protein of modest molecular size, whereas chitin synthase is membrane-bound and substantially larger. In contrast to insect chitinases, which have been studied extensively, no insect chitin synthase has been well characterized. Recently, however, Tellam et al. (2000) reported that the cDNA for a chitin synthase of the blow fly, *Lucilia cuprina*, encodes a large transmembrane protein containing 1592 amino acids. Subsequently Ibrahim et al. (2000) described a cDNA for the same enzyme from the mosquito, *Aedes aegypti*, which encoded a protein only about half the size of the blow fly enzyme, containing only 865 amino acids.

We have recently determined the sequence of a full-length cDNA for a chitin synthase-like protein from the tobacco hornworm, *Manduca sexta*. This was initiated by using degenerate primers designed from conserved regions of chitin synthase protein sequences of fungi and nematodes for use in polymerase chain reaction (PCR) amplification from hornworm cDNA and genomic DNA. The *Manduca* chitin synthase cDNA (*MsCHS1*) has an open reading frame of 4692 nucleotides that encodes a 1564-amino acid transmembrane protein. The hornworm enzyme is very much like the blow fly and fruit fly enzymes in sequence, domain structure, and predicted transmembrane profile. Transcription studies were undertaken to determine whether *MsCHS1* is expressed during cuticle deposition.

## 2. Materials and methods

### 2.1. Insect cultures

*M. sexta* larvae were reared using an artificial diet at 27°C as described by Bell and Joachim (1976). The photoperiod was 16 h of light and 8 h of darkness.

### 2.2. Genomic DNA and cDNA sequencing

Genomic DNA was extracted from third and fourth instar larvae that had been homogenized in 100 mM

NaCl, 10 mM Tris-HCl (pH 8.0), 25 mM EDTA, 0.5% sodium dodecyl sulfate, and phenol/chloroform/isoamyl alcohol (25:5:1, v/v/v) (Ausubel et al., 1994). A chitin synthase DNA fragment was amplified by PCR from cDNA prepared from fifth instar larval mRNA. The degenerate primers Chs1F (5'-CARAANTTYGAR-TAYRMNAT-3', where R=A or G; N=G, A, T or C; Y=C or T and M=A or C, corresponding to positions 2476–2495 in Fig. 2) and Chs2R (5'-CCANCKNCKNCKYTGNBHDWWRAAYTC-3', where K=G or T; B=T, C or G; H=A, C or T; D=A, G or T and W=A or T, corresponding to positions 2761–2787, see Figs. 1 and 2) were designed from the conserved regions Q-K/N-FEY-A/K-I/M and EF-Y/F/I-N/S/G-QRRRW found in chitin synthases of fungi and nematodes [Din and Yarden, 1994; Yanai et al., 1994; Nino-Vega et al., 1998; Bradshaw, 1999 (unpublished, GenBank Protein Database Accession No. T32452); McMurray, 1999 (unpublished, GenBank Protein Database Accession No. T25284)]. PCR samples (50 µl) contained 10 mM Tris-HCl at pH 9, 1.5 mM MgCl<sub>2</sub>, 0.5 µM of each primer, 50 mM KCl, 0.1 mM of each dNTP, 0.05 unit/µl of Taq DNA polymerase, and approximately 100 ng of *M. sexta* DNA template, and reactions were performed using a PTC-100 thermocycler (MJ Research, Inc., Watertown, MA). DNA was initially denatured for 3 min at 94°C, and the PCR amplification was conducted for 45 cycles, with 30 s denaturing at 94°C, 30 s annealing at 58°C, 1 min extension at 72°C, and a final 10 min extension at 72°C. The PCR product was subjected to electrophoresis on 1% low melting point agarose gel, and the resulting DNA fragment (~300 bp) was excised and purified using the Wizard PCR Preps DNA purification kit (Promega, Madison, WI). The PCR fragment was cloned into a pGEM-T vector (Promega), and its sequence was determined using an automated sequencer (ABI Prism 3700, DNA Sequence and Synthesis Facility, Iowa State University, Ames, IA).

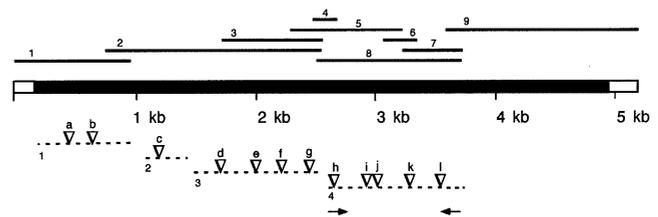


Fig. 1. Cloning strategies for obtaining the full cDNA sequence of *M. sexta* chitin synthase. Solid and dashed lines represent cDNA and genomic DNA fragments, respectively. The former fragments were sequenced in both directions. Hollow bars are 5'- and 3'-untranslated regions, and solid bar is the open reading frame. cDNA fragment 1 was RT cDNA obtained from 5'-RACE. cDNA fragments 2–9 were amplified from RT cDNA and cDNA libraries. Genomic DNA fragments 1–4 include 12 introns (a–l) that are 213, 328, 686, 471, 83, 86, 311, 76, 464, 76, 78, and 82 bp in length, respectively. Arrows below genomic DNA fragment 4 identify the position of primers used for RT-PCR analysis.

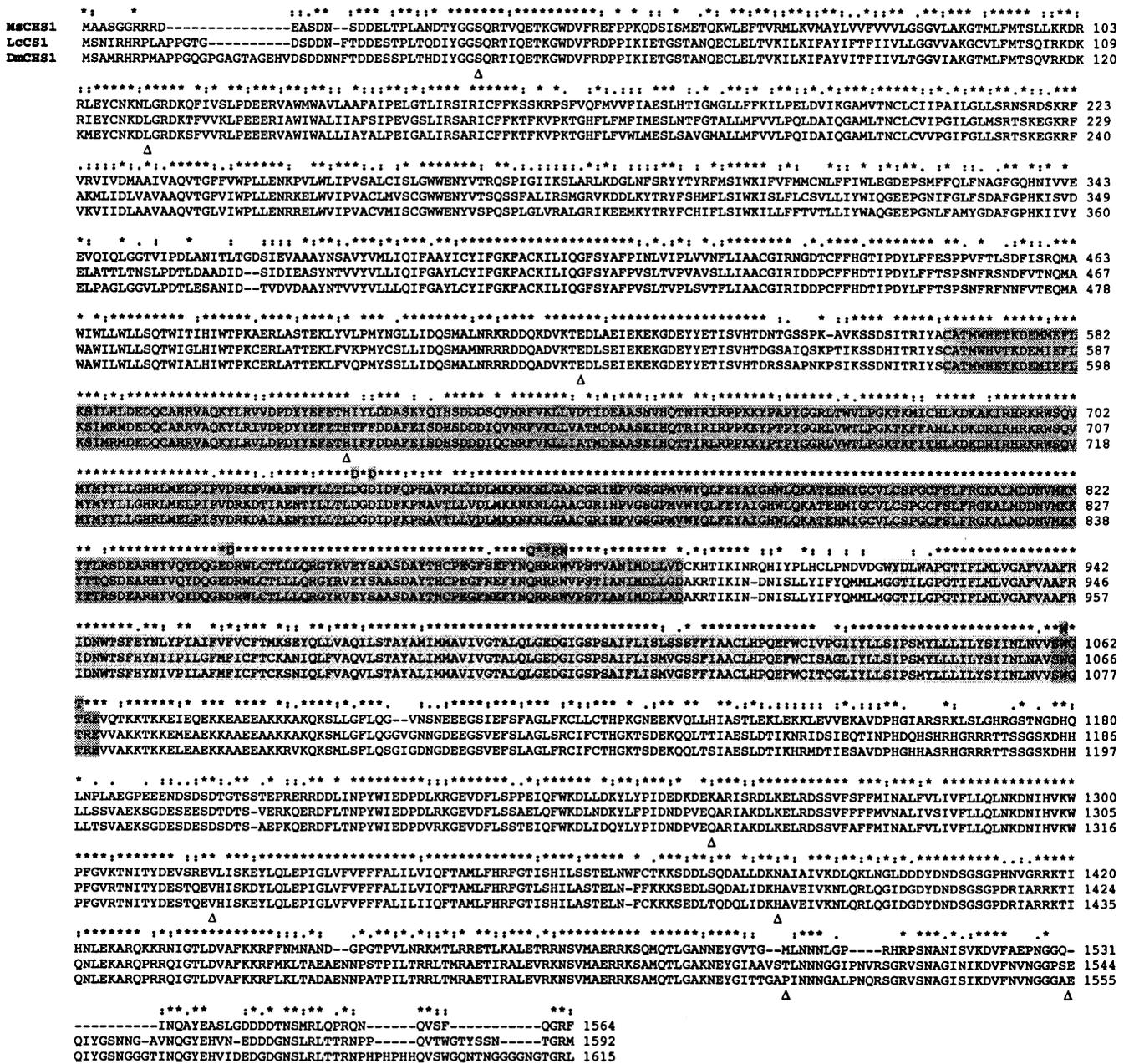


Fig. 2. Alignment of predicted protein sequences of *Manduca* (MsCHS1), *Lucilia* (LcCS-1) and *Drosophila* (DmCHS1) chitin synthases using ClustalW. Symbols above each set of lines indicate identity (\*), highly conserved substitutions (:), conserved substitutions (.) or amino acids believed to be important for catalysis. Shaded lines are believed to encompass the catalytic domain (darkly shaded) and the transmembrane region for the chitin polymer to traverse the plasma membrane (lightly shaded). The relative positions of introns in the genomic DNA sequence that encodes DmCHS1 (GenBank Accession: AE003603) are indicated by delta (Δ). All introns we found use conventional 5'- and 3'-splice junctions and were located to maximize protein sequence identity with MsCHS1 predicted from the cDNA clone.

2.3. cDNA cloning

To obtain the sequence of the full-length cDNA of the *Manduca* CHS1 gene, overlapping cDNA fragments were assembled. RNA was isolated from different stages of *Manduca* development (pre-molting, actively feeding, and post-molting), different instars (first–fifth larval instars), and from prepupae and newly ecdysed pupae. Total RNA was prepared with TRIzol reagent (Life-

Tech, Rockville, MD) and mRNA was purified with a Messenger RNA Isolation Kit (Stratagene, La Jolla, CA). Reverse transcriptions and subsequent PCR amplifications were performed by using 5'- and 3'-RACE kits (Life-Tech, Rockville, MD). Several chitin synthase cDNA fragments of *MsCHS1* were cloned from these PCR fragments by screening a fifth instar larval cDNA library using fragment 4 as the probe and sequenced (Fig. 1). RNA that was prepared from third instar or

younger larvae was found to be the best template for preparation of cDNA containing the 5'- and 3'-ends of the gene. Each fragment was sequenced in both directions and overlapped a cloned, contiguous fragment by more than 200 bp. The continuity of the sequence in the assembled full-length cDNA was also verified by sequencing genomic DNA that had been amplified by PCR. DNA sequences were generated by automated sequencing as described above.

The cDNA sequence of the *Manduca* chitin synthase, *MsCHS1*, was subjected to homology searches (BLASTX) of the GenBank database provided by the National Center for Biotechnology Information (Altschul et al., 1997; Gish and States, 1993). The Wisconsin Sequence Analysis Package (GCG Unix version 10, Genetics Computer Group., Madison, WI) and DNA-STAR Software package (Madison, WI) were used to compare chitin synthase protein sequences. Sequence analysis tools of the ExPASy Molecular Biology Server of the Swiss Institute of Bioinformatics were also used to identify features of deduced protein sequences.

#### 2.4. Developmental expression and northern blotting

Total RNA was prepared from the epidermis of fifth instar larvae or pupae using the method of Chomczynski and Sacchi (1987). Three to six insects were dissected for each data point. To facilitate the homogenization of the tissue, each sample was frozen in liquid nitrogen and then ground to a powder with a mortar and pestle prior to RNA isolation. Two micrograms of total RNA was reverse transcribed with SUPERScript II RNase H-Reverse Transcriptase (Life Technologies, Rockville, MD) in a total volume of 20  $\mu$ l, according to the manufacturer's directions. One microliter of each reverse transcribed reaction was utilized for PCR under the following conditions: initial denaturation at 94°C for 2 min; for each cycle, denaturation was at 94°C for 30 s, annealing at 56°C for 30 s, and extension at 72°C for 2 min. Primers used were: (2507) 5'-TGCAAAGGCGACTGAACACA-3' (2527) and (3714) 5'-TTCCTCTGGTCCCTCAGCTA-3' (3695) for *MsCHS1*. Primers used for ribosomal protein S3 were (221) 5'-CGCGAGTTGACTTCGGT-3' (237) and (619) 5'-GCCGTTCTTGCCCTGTT-3' (603). Amplification of chitin synthase cDNA was for 28 cycles and that of ribosomal protein S3 cDNA for 25 cycles. One-fifth of each PCR sample was fractionated in a 1% agarose gel in TAE buffer, and the amplified fragments were visualized by staining with ethidium bromide. Gels were scanned and analyzed with the Kodak Digital Science ID imaging software (version 2.0.3).

For northern blot analysis, 20  $\mu$ g of total RNA from pupal epidermis day-0 (P-0) was fractionated in a 1% agarose/1.2% formaldehyde gel, and transferred to a nitrocellulose membrane. The probe for hybridization

was a 1.2-kb cDNA radiolabeled by the Multiprime DNA Labeling System (Amersham Pharmacia Biotech). Hybridization was performed under high stringency conditions (5X SSC/5X Denhardt's/50% formamide (v/v)/0.1% SDS/100  $\mu$ g ml<sup>-1</sup> denatured salmon sperm DNA) at 42°C. Following hybridization, the membrane was washed twice in 2X SSC/0.1% SDS at room temperature, twice in 2X SSC/0.1% SDS at 60°C, and once in 0.2X SSC/0.1% SDS at 60°C; each wash was for 15 min. The membrane was then subjected to autoradiography with an intensifying screen at -70°C.

### 3. Results and discussion

Chitin synthase genes of invertebrates have been identified from only four species of insects and only two of nematodes, while being greatly surpassed in number by those identified from fungi. Even though chitin plays such an important role in insect development and physiology, it is still unclear how many chitin synthases a species may have or how they are regulated. We have begun to address some of these questions by characterizing a putative chitin synthase gene (*MsCHS1*) found in *M. sexta*, the tobacco hornworm.

A fragment containing a portion of the *MsCHS1* sequence was identified initially by PCR using cDNA prepared from fifth instar larval mRNA as the template and primers corresponding to two highly conserved regions in fungal chitin synthases. That amplified fragment of 312 bp (positions 2476–2787 in Fig. 1, see fragment #4 at top) had sequence similarity by BLASTX searches not only to invertebrate and fungal chitin synthases but also to other family 2 glycosyltransferases, notably rhizobial *N*-acetylglucosaminyl transferase, NodC, and the hyaluronan synthases (Campbell et al., 1997). Family GT-2 glycosyltransferases, which has more than 950 members, synthesize a wide spectrum of products ranging from chitin, cellulose, and hyaluronan to complex cell surface glycolipids (Coutinho and Henrissat, 1999; Henrissat and Davies, 2000; Rini and Sharon, 2000).

A full-length cDNA sequence was then determined by assembly of overlapping sequences of cDNA clones isolated from fifth instar larval cDNA libraries, and from PCR fragments generated from third instar larval cDNA by 5'- and 3'-RACE. The assembled full-length cDNA was confirmed to be transcribed from a unique locus by sequencing of several genomic DNA clones covering this region using several primers (Fig. 1). The cDNA of *MsCHS1* contained 5152 nucleotides with an ORF that spans from nucleotide position 145 to 4836 (Fig. 1). A polyadenylation signal, AATAAA, was found at positions 5118–5123. The sequence of *MsCHS1* has been deposited with GenBank under Accession No. AY062175.

The *MsCHS1* ORF was deduced to encode a protein of 1564 amino acids (Fig. 2) with a molecular mass of approximately 179 kDa and a calculated pI value of pH 7.04. Chitin synthases of other insect species, specifically one known from the sequence of a cDNA of *Lucilia LcCS-1* (1592 aa, GenBank Accession No. AAG09712) and that deduced by us from genomic DNA for *Drosophila DmCHS1* (1615 aa, GenBank Accession No. AE003603) are comparable in size.

Sequence alignments generated from BLAST and CLUSTALW (1.81) analyses indicated that the protein sequence of *MsCHS1* was also most similar over its entire length to the *Lucilia LcCS-1* and *Drosophila DmCHS1* sequences. All three of these sequences are aligned in Fig. 2. Pairwise comparisons estimated sequence identities of 68% and sequence similarities of 80% between *MsCHS1* and either *LcCS-1* or *DmCHS1*. Other chitin synthase-like proteins from invertebrates had less identity to *MsCHS1* and the sequences could not be aligned in full (data not shown). These included three from insect species (*Drosophila DmCHS2*, GenBank Accession No. AAF51798; *Aedes AaCHS1*, GenBank Accession No. AAF34699; and *Anopheles AgCHS1*, GenBank Accession No. AY056833) as well as three from nematodes (*Caenorhabditis CeCHS1*, GenBank Accession No. T32452; *CeCHS2*, GenBank Accession No. T25284; and *Brugia Bm-chs-1* GenBank Accession No. AF274311; Harris et al., 2000). The region that had the highest concentration of sequence identities was demarcated by the common block sequences, CATMWHXT (residues 567–574 of *MsCHS1*) and WGTRE (residues 1061–1065). Also, within this region are the blocks QXFEY and QRRRW used to design the PCR primers as well as several key aspartic acids (D) believed to be important for catalysis of family 2 glycosyltransferases (Breton et al., 1998; Wiggins and Munro, 1998; Breton and Imberty, 1999).

To provide some additional perspective to the comparison of primary sequences, secondary structure predictions were made for *MsCHS1* and compared. Yeast chitin synthases are localized to the plasma membrane with a signal anchor or uncleaved signal peptide and are predicted to have multiple transmembrane helices for retention and topology. When analyzed by PSORT and SignalP V2.0 software, the *M. sexta* chitin synthase was predicted to be a non-secreted, plasma membrane protein ([www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/); Nielsen et al., 1997; Nielsen and Krogh, 1998). Likewise, secondary structure prediction using TMHMM (v.2.0) software ([www.cbs.dtu.dk/services/](http://www.cbs.dtu.dk/services/); Hirokawa et al., 1998; Sonnhhammer et al., 1998) indicated that *MsCHS1* was a membrane-associated protein. The tobacco hornworm enzyme apparently has 16 transmembrane helical regions that are from 19 to 23 amino acids long, as depicted in Fig. 3. These regions are residues 73–95, 128–150, 162–184, 194–216, 223–245, 250–267, 300–319, 370–392,

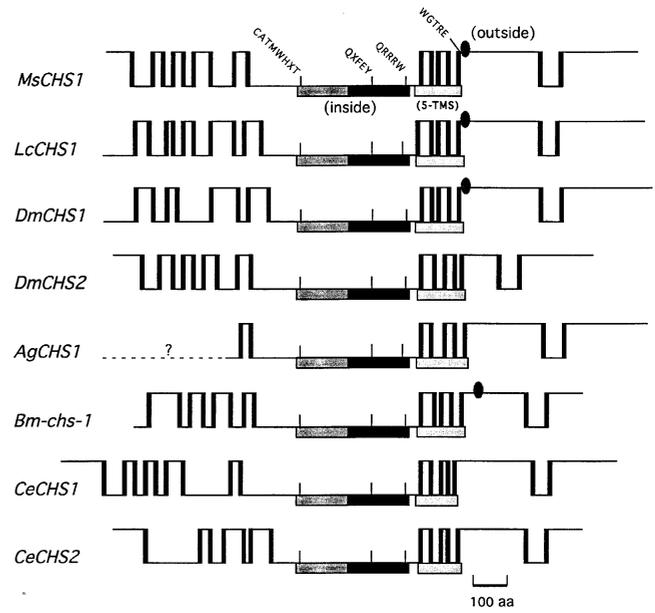


Fig. 3. Predicted transmembrane profiles of putative invertebrate chitin synthases using TMHMM (v. 2.0). Protein sequences analyzed were *MsCHS1*, *LcCS-1*, *DmCHS1*, *Drosophila DmCHS2* (GenBank AE003596), *Anopheles AgCHS1* (AY056833), *Brugia Bm-chs-1* (AF274311), and *Caenorhabditis CeCHS1* (AF026210) and *CeCHS2* (Z72516). Transmembrane helices are indicated as vertical bars and regions predicted to be on the inside or outside of the plasma membrane by horizontal lines. For orientation, blocks of highly conserved amino acids are indicated. The probable catalytic domain is boxed; the darker shaded portion depicts the region with high sequence similarity to fungal chitin synthases and the lighter shaded portion with low sequence similarity to fungal chitin synthases. The transmembrane spans that may be important for extrusion of chitin through the plasma membrane are indicated as 5-TMS. Shaded ovals identify coiled-coil domains (Paircoil Program).

405–427, 920–942, 949–968, 973–995, 1002–1024, 1034–1056, 1272–1291, and 1329–1351.

Also notable in the amino acid sequence of *MsCHS1* are three long non-transmembrane regions from residues 430–919, 1059–1269 and 1352–1564. The first of these non-transmembrane regions was predicted to be oriented to the cytoplasm, whereas the other two are likely to face the extracellular space. Within the 430–922 region, there is a block of sequence that is highly conserved among all chitin synthases (dark shaded box in Fig. 3). Alterations by site-directed mutagenesis of some of these highly conserved amino acids (indicated in Fig. 2) of the *Saccharomyces* chitin synthases CHS2 and CHS3 sequences (Nagahashi et al., 1995; Yabe et al., 1998; Cos et al., 1998) strongly suggest that this region is the catalytic domain (shown in greater detail in Fig. 4). This portion of the catalytic domain is believed to be oriented to the cytoplasm given that one of the substrates for chitin synthesis, UDP-GlcNAc, is found in the cytoplasm.

The transmembrane profile of *MsCHS1* was compared to those of putative chitin synthases from insects and nematodes (Fig. 3). A consistent pattern can be seen

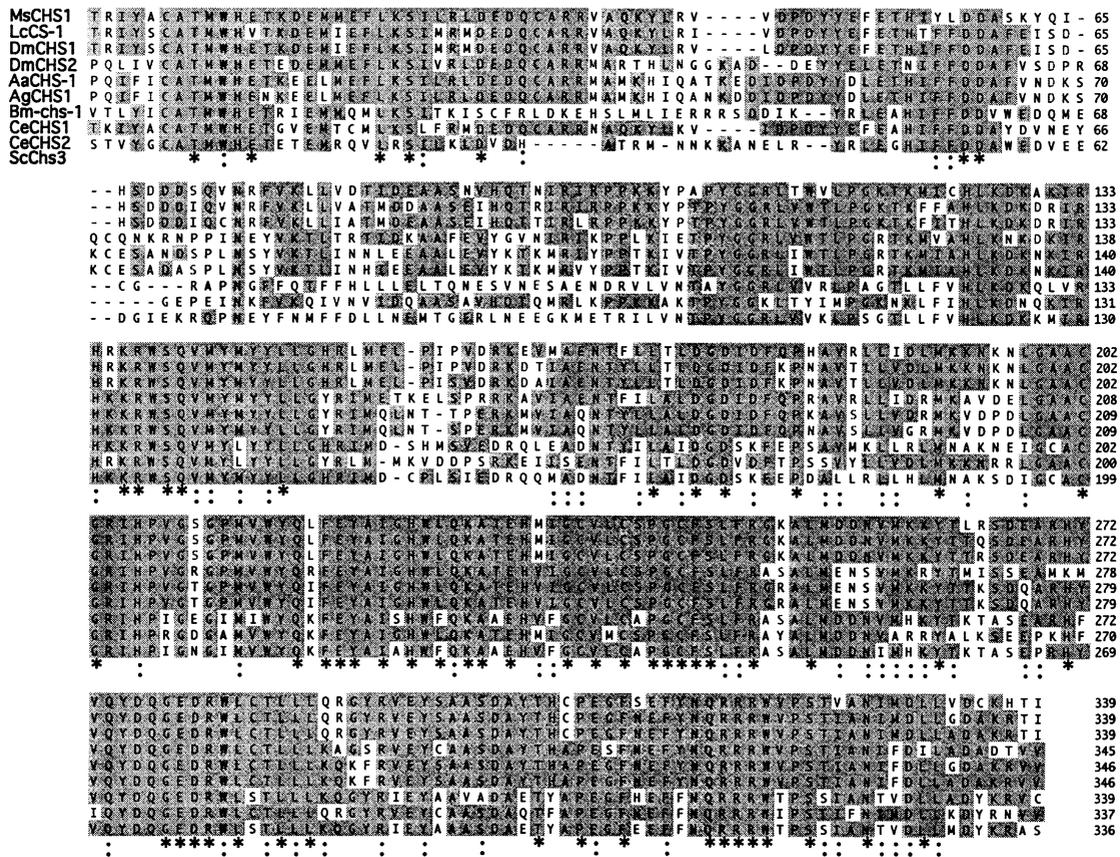


Fig. 4. Alignment of predicted protein sequences within the putative catalytic domain of invertebrate chitin synthases. Sequences were aligned using MegAlign (DNASTar, Inc.) running the Clustal (PAM250) program. Abbreviations to designate proteins are the same as in Fig. 3 with the sequence of *Aedes* AaCHS1 (GenBank AAF34699) added. For the representative fungal chitin synthase sequence of *Saccharomyces* ScCHS3 (P29465), only the amino acids identical (\*) or highly conserved (:) with the majority of the invertebrate sequences are indicated. Shaded amino acids identify amino acids identical among three or more sequences.

when topology was predicated on the highly conserved catalytic domain being oriented to the cytoplasm (inside). A feature of particular interest is the predicted five transmembrane spans (5-TMS) that immediately follow the catalytic domain. These spans can also be found in the same relative position with respect to the catalytic domain in fungal chitin synthases (data not shown). Even more significant is the fact that other glycosyl-transferases that synthesize extracellular polymers such as cellulose also have this proximate feature consisting of a catalytic domain next to a potential pore-forming structure; the 5-TMS may contribute to a pore in the membrane through which the newly synthesized carbohydrate polymers are extruded, as was proposed by Richmond (2000) for cellulose synthases.

When the entire *A. aegypti* cDNA (GenBank AF223577) was used as a query of the GenBank protein database by BLASTX, we found that there were significant sequence similarities between 5'-upstream sequences and 3'-downstream sequences in chitin synthases of LcCS-1 and MsCHS1, which were not found in the AaCHS1 ORF identified by Ibrahim et al. (2000). We suggest that the *Aedes* protein is actually longer than

the 865 amino acids reported and that the *Aedes* ORF probably extends further in both the 5'- and 3'-directions. Thus, the *Aedes* CHS1 cDNA sequence needs to be re-evaluated. The *Aedes* chitin synthase protein sequence listed in the database has a transmembrane profile that also suggests the amino acid sequence to be truncated (data not shown). Prediction of transmembrane helices revealed that the *Aedes* protein is missing the first seven putative transmembrane helices as well as the last one present in the other chitin synthases (Fig. 3). Similarly, the sequence of *Anopheles* AgCHS1 may also be missing the first seven putative transmembrane helices and may represent a partial sequence. However, both the *Aedes* and *Anopheles* sequences do contain the putative catalytic domain and also exhibit the 5-TMS helices that follow on the C-terminal side.

Four of the invertebrate enzymes, MsCHS1, DmCHS1, LcCS-1 and Bm-chs-1, were also predicted by the Paircoil program (Berger et al., 1995) to have a coiled-coil domain (small oval in Fig. 3) that immediately follows the 5-TMS and the conserved sequence WGTRE. The coiled coil predicted for Bm-chs-1 is slightly more distant from WGTRE than for the other

three proteins. This coiled-coil domain may serve as a site for a protein–protein interaction and, given its potential binding location close to the catalytic domain, may help to regulate chitin synthesis.

Next, we focused our attention on the region implicated as important for catalytic activity of chitin synthases. Initial comparisons of the full-length protein sequence of MsCHS1 showed that it, as well as other invertebrate chitin synthase sequences, had similarity to the Pfam domain Chitin—synth—2 (Accession No. PF03142), represented by Class IV and V fungal chitin synthases (Specht et al., 1996). Reiterated BLASTP revealed that the similarities to the fungal chitin synthases could be extended approximately 500 residues from about CATMWHXT to WGTRE, as identified in Fig. 3. Among fungal chitin synthases, regions at the N- and C-termini outside of this central conserved region are quite variable. Within this interior region there was also similarity to a portion of the glycosyltransferase domain, Pfam PF00535, that represents a broad spectrum of proteins that include *N*-acetylglucosaminyl transferases (Unligil et al., 2000). To estimate the similarities of insect and nematode chitin synthases with those of fungi, the comparable span of amino acid sequences was compared and included the signature motifs CATMWHXT and QRRRW of the invertebrate and fungal sequences. Blocks of identical/conserved sequences in the 5-TMS to WGTRE region were accompanied by an increased number of gaps, and, therefore, this aligned region was not included in the comparison. Results of CLUSTAL (PAM250) analyses of the 10 insect, nematode and fungal catalytic domains are shown in Fig. 4, and the sequence identities and similarities are summarized in Table 1. *Saccharomyces* Chs3p was chosen as representative of the fungal chitin synthases that were used to generate Pfam domain PF03142 because it is the best characterized of the fungal enzymes. Other fungal chitin synthases from *Aspergillus fumigatus* (Mellado et al., 1995), *Emericella nidulans* (Fujiwara et al., 1997), and *Blumeria graminis* (Zhang et al., 2000) exhibit overall

identities of 35, 36 and 36%, respectively, with MsCHS1.

There are several conserved motifs present in the putative catalytic domain of MsCHS1 and the other invertebrate chitin synthases, as well as the fungal enzymes. These include the DxD motif found in almost all classes of glycosyltransferases and, as identified by the glycosyltransferase domain Pfam 00535, the invertebrate sequences have DGD (positions 736–738) as indicated in both Figs. 2 and 4. The aspartic acid (D) residues are hypothesized to play a major role in enzyme function as a part of either the catalytic site or a binding site for a nucleotide sugar or metal ion (Breton et al., 1998; Wiggins and Munro, 1998). The invertebrate chitin synthases also have motifs that have been identified as important for catalytic activity of fungal chitin synthases based on site-directed mutagenesis of *S. cerevisiae* chitin synthases ScCHS2 and ScCHS3 (Nagahashi et al., 1995; Cos et al., 1998). In addition to mutagenesis of the DxD motif, results obtained from mutagenesis of the motifs EDR and QRRRW in the fungal enzymes revealed that they also contain key amino acids of the catalytic site (see Figs. 2 and 4). Another motif implicated to play a role in catalysis (synthesizing polymeric products) by site-directed mutagenesis studies (Yabe et al., 1998) and conserved in fungal chitin synthases is SWGTRE (identified in Figs. 2 and 3). The similarities in amino acid sequences between putative *M. sexta* and other invertebrate chitin synthases to those microbial proteins with chitin synthase activity strongly support the hypothesis that the invertebrate proteins are actually chitin synthases.

With completion of the *Drosophila* Genome project, it has become evident that there are likely only two genes that encode chitin synthases or related proteins in the fruit fly (*DmCHS1* and *DmCHS2* of Figs. 2–4). The alignment of Fig. 4 was used to generate a tree that finds *DmCHS1* and *DmCHS2* in two branches of the insect chitin synthases as shown in Fig. 5. So far as can be surmised from available sequences, the insect chitin syn-

Table 1  
Similarities of sequences from the catalytic region of chitin synthases<sup>a</sup>: % identity/(% similarity)

|          | MsCHS1 | LeCS-1 | DmCHS1 | DmCHS2 | AgCHS1 | AaCHS1 | Bm-chs-1 | CeCHS1 | CeCHS2 | ScCHS3 |
|----------|--------|--------|--------|--------|--------|--------|----------|--------|--------|--------|
| MsCHS1   |        | 87     | 87     | 69     | 70     | 70     | 52       | 69     | 54     | 15     |
| LeCS-1   | (92)   |        | 95     | 67     | 70     | 70     | 52       | 68     | 54     | 18     |
| DmCHS1   | (92)   | (96)   |        | 68     | 72     | 71     | 53       | 69     | 55     | 18     |
| DmCHS2   | (79)   | (78)   | (79)   |        | 74     | 74     | 53       | 61     | 56     | 15     |
| AgCHS1   | (84)   | (85)   | (85)   | (84)   |        | 96     | 53       | 60     | 53     | 16     |
| AaCHS1   | (84)   | (86)   | (85)   | (84)   | (97)   |        | 52       | 60     | 54     | 16     |
| Bm-chs-1 | (66)   | (65)   | (66)   | (67)   | (69)   | (67)   |          | 51     | 73     | 16     |
| CeCHS1   | (83)   | (82)   | (82)   | (74)   | (75)   | (76)   | (66)     |        | 51     | 15     |
| CeCHS2   | (68)   | (69)   | (70)   | (69)   | (66)   | (66)   | (83)     | (69)   |        | 17     |
| ScCHS3   | (25)   | (30)   | (30)   | (23)   | (25)   | (24)   | (24)     | (24)   | (25)   |        |

<sup>a</sup> Each protein sequence used in these pair-wise comparisons was the same as those shown in Fig. 4.

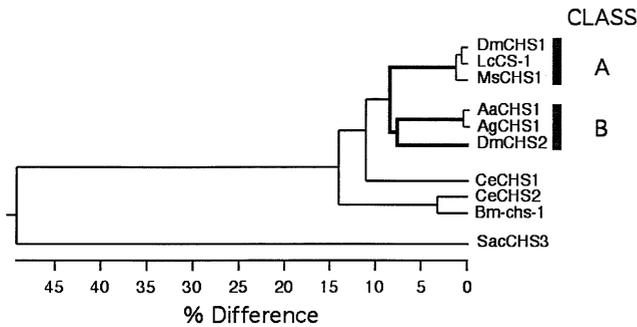


Fig. 5. Classification of the insect chitin synthases. The tree was generated by MegAlign (DNASTAR, Inc.) from the alignment shown in Fig. 4. Sequence abbreviations are the same as in Figs. 3 and 4.

these family can be divided into two classes following the scheme developed by Bowen et al. (1992) for the multiple chitin synthases of fungi. *MsCHS1* is in Class A with *DmCHS1* and *LcCS-1*, whereas in Class B are *DmCHS2*, *AaCHS1* and *AgCHS1*. Consistent with this division is the confinement of expression of *LcCS-1* in epidermal cells for synthesis of chitin in the cuticle (Tellam et al., 2000), and of *AaCHS1* in epithelial cells of the midgut for synthesis of chitin in the peritrophic matrix (Ibrahim et al., 2000).

Lastly, we examined the expression of *MsCHS1* to determine whether it was associated with cuticle formation, as was suggested by its placement in the Class A enzymes' branch (Fig. 5). The transcript of *MsCHS1* was

detected by northern blot analysis and by RT-PCR. Northern blot analysis detected a major transcript of approximately 5.3 kb (Fig. 6A), indicating that our cDNA is full length or nearly so. Upon longer exposure, a minor transcript of approximately 7.3 kb was also detected. This larger transcript may represent pre-mRNA or the usage of alternative promoters or polyadenylation sites, or may be the product of a closely related gene, i.e. one for a Class B chitin synthase. For RT-PCR of *MsCHS1* transcript, primers were selected that would amplify a 1.2 kb product from mRNA and a 2.0 kb product from genomic DNA (see Fig. 1); the latter product was not detected, indicating the absence of genomic DNA contamination. RT-PCR showed chitin synthase transcript levels to be relatively constant in the epidermis during the feeding stage of the fifth instar (Fig. 6B). Transcript levels then dropped to near undetectable levels after feeding ceased and gradually increased during the wandering prepupal stage as the insect approached the pupal molt. To verify that the reduction in transcript levels during the wandering stage was not due to a poor quality of RNA and its inability to function as a template or to lower amounts of template RNA for the reverse transcription reaction, RT-PCR was also used to examine the expression of ribosomal protein S3 (*rpS3*), a constitutively expressed housekeeping gene (Jiang et al., 1996). The transcript levels for *rpS3* remained relatively constant for all time points tested, even during the wandering stage. These results are consistent with the need for chitin synthesis to occur in the epidermis during both the growth and molting periods of development (Cohen, 1991, 2001). They also suggest a negative regulation of the CHS gene by ecdysteroid because hormonal levels increase just prior to wandering behavior and decline before pupation (Baker et al., 1987).

In summary, in insects and other organisms, the precursor molecule UDP-N-acetylglucosamine is polymerized by chitin synthases attached to the plasma membrane to make extracellular chitin (Berninsone and Hirschberg, 2000). From the sequences of the *Drosophila* genome, the projected number of chitin synthase genes used by this insect is two. Confirmation of the two predicted *Drosophila* protein sequences awaits completion and analysis of full-length cDNA sequences. With a full-length cDNA for the *Manduca* chitin synthase gene, further study of chitin synthesis in insects, which was not possible previously because the membrane-bound native enzyme could not be isolated and characterized, will be possible. Future experiments include recombinant expression of the full-length protein and also just the catalytic domain as an active soluble form to be used for elucidating the mechanism of action (Machida et al., 2001), antibody production, tissue and cellular localizations, and hormonal regulation of *MsCHS1* gene expression.

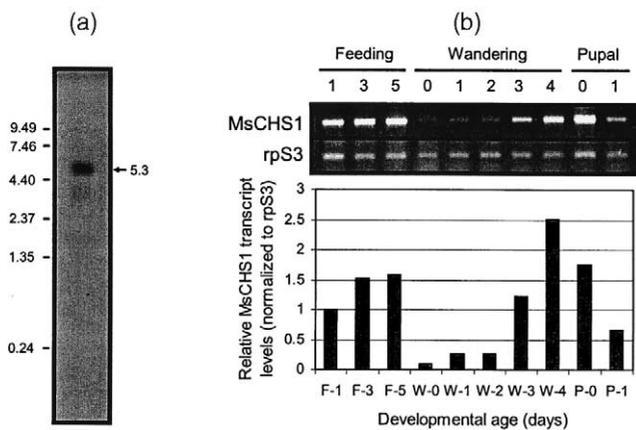


Fig. 6. Analysis of *MsCHS1* gene expression in epidermis. (A) Analysis by northern blot hybridization. Twenty micrograms of total RNA prepared from day-0 pupal epidermis (P-0) was fractionated by gel electrophoresis. The size of the *MsCHS1* transcript in kilobases is indicated by the arrow on the right. Standards in kilobases are shown on the left. (B) Analysis by RT-PCR. PCR was performed with cDNA synthesized from total RNA prepared from the epidermis of fifth instar larvae or pupae. The age, in days, for each developmental stage is noted above each lane. Relative amounts of *MsCHS1* transcript (1188 bp) to that of *rpS3* (398 bp) were determined for each stage by densitometry and graphed. Abbreviations are F, feeding stage; P, pupal stage; *rpS3*, ribosomal protein S3; W, wandering stage.

## Acknowledgements

We thank Yimin Lu for one of the *M. sexta* cDNA libraries, Dr J. Fuhrman for nematode CHS gene sequence information, and Drs Fuhrman, P.W. Robbins and A. Retnakaran for reviewing the manuscript. This research was supported in part by a USDA-NRI grant to SMK and KJK, an NSF grant to MRK and KJK, and the Kansas Agricultural Experiment Station (Contribution No. 02-385-J). CAS acknowledges support from NIH grants GM 31318 to P.W. Robbins and RO1 A125780 to S.M. Levitz. All programs and services of the USDA are offered on a nondiscriminatory basis without regard to race, color, national origin, religion, sex, age, marital status or handicap.

## References

- Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Venter, J.C., 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 1994. *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc.
- Baker, F.C., Tsai, L.W., Reuter, C.C., Schooley, D.A., 1987. In vivo fluctuations of JH, JH acid, and ecdysteroid titer, and JH esterase activity, during development of fifth stadium *Manduca sexta*. *Insect Biochem.* 17, 989–996.
- Bell, R.A., Joachim, F.G., 1976. Techniques for rearing laboratory colonies of tobacco hornworms and pink bollworms. *Ann. Entomol. Soc. Am.* 69, 365–373.
- Berger, B., Wilson, D.B., Wolf, E., Tonchev, T., Milla, M., Kim, P.S., 1995. Predicting coiled coils by use of pairwise residue correlations. *Proc. Natl. Acad. Sci. U S A* 92, 8259–8263.
- Berninson, P.M., Hirschberg, C.B., 2000. Nucleotide sugar transporters of the tiolgi apparatus. *Curr. Opin. Struc. Biol.* 10, 542–547.
- Bowen, A.R., Chen-Wu, J.L., Momany, M., Young, R., Szanislo, P.J., Robbins, P.W., 1992. Classification of fungal chitin synthases. *Proc. Natl. Acad. Sci. U S A* 89, 519–523.
- Breton, C.L., Bettler, E., Joziassé, D.H., Geremia, R., Imberty, A., 1998. Sequence function relationships of prokaryotic and eukaryotic galactosyltransferases. *J. Biochem.* 123, 1000–1009.
- Breton, C.L., Imberty, A., 1999. Structure/function studies of glycosyltransferases. *Curr. Opin. Struc. Biol.* 9, 563–571.
- Bulawa, C.E., 1993. Genetics and molecular biology of chitin synthesis in fungi. *Annu. Rev. Microbiol.* 47, 505–534.
- Cabib, E., Roh, D., Schmidt, M., Crotti, L.B., Varma, A., 2001. The yeast cell wall and septum as paradigms of cell growth and morphogenesis. *J. Biol. Chem.* 276, 19679–19682.
- Campbell, J.A., Davies, G.J., Bulone, V., Henrissat, B., 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* 326, 929–939 (Correction: 1998 *Biochem. J.* 329, 719).
- Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 156–159.
- Cohen, E., 1991. Chitin biochemistry. In: Binnington, K., Retnakaran, A. (Eds.), *Physiology of the Insect Epidermis*. CSIRO Publishers, Melbourne, pp. 94–112.
- Cohen, E., 2001. Chitin synthesis and inhibition: a revisit. *Pest Management Sci.* 57, 946–950.
- Coutinho, P.M., Henrissat, B., 1999. Carbohydrate-active enzymes: an integrated database approach. In: Gilbert, H.H., Davies, G.J., Henrissat, H., Svensson, B. (Eds.), *Recent Advances in Carbohydrate Bioengineering*, Carbohydrate-active enzymes family server URL: <http://afmb.cnrsmrs.fr/~pedro/CAZY/db.html>. Royal Society of Chemistry, Cambridge, UK, pp. 3–12.
- Cos, T., Ford, R.A., Trilla, J.A., Duran, A., Cabib, E., Roncero, C., 1998. Molecular analysis of Chs3p participation in chitin synthase III activity. *Eur. J. Biochem.* 256, 419–426.
- Din, A.B., Yarden, O., 1994. The *Neurospora crassa* chs-2 gene encodes a non-essential chitin synthase. *Microbiology* 140, 2189–2197.
- Fujiwara, M., Horiuchi, H., Ohta, A., Takagi, M., 1997. A novel fungal gene encoding chitin synthase with a myosin motor-like domain. *Biochem. Biophys. Res. Commun.* 236, 75–78.
- Gish, W., States, D.J., 1993. Identification of protein coding regions by database similarity search. *Nat. Genet.* 3, 266–272.
- Harris, M.T., Lai, K., Arnold, K., Martinez, H.F., Specht, C.A., Fuhrman, J.A., 2000. Chitin synthase in the filarial parasite, *Brugia malayi*. *Mol. Biochem. Parasitol.* 111, 351–362.
- Henrissat, B., Davies, G.J., 2000. Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. *Plant Physiol.* 124, 1515–1519.
- Hirokawa, T., Boon-Chieng, S., Mitaku, S., 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14, 378–379.
- Ibrahim, G.H., Smartt, C.T., Kiley, L.M., Christensen, B.M., 2000. Cloning and characterization of a chitin synthase cDNA from the mosquito *Aedes aegypti*. *Insect Biochem. Mol. Biol.* 30, 1213–1222.
- Jiang, H., Wang, Y., Kanost, M., 1996. Primary structure of ribosomal proteins S3 and S7 from *Manduca sexta*. *Insect Mol. Biol.* 5, 31–38.
- Koga, D., Mitsutomi, M., Kono, M., Matsumiya, M., 1999. Biochemistry of chitinases. In: Jollés, P., Muzarelli, R.A.A. (Eds.), *Chitin and Chitinases*. Birkhäuser-Verlag, Basel, Switzerland, pp. 111–123.
- Kramer, K.J., Koga, D., 1986. Insect chitin: physical state, synthesis, degradation and metabolic regulation. *Insect Biochem* 16, 851–877.
- Machida, S., Niimi, S., Shi, X., 2001. Expression of the cytoplasmic domain of NodC as an active form in *Drosophila* S2 cells. *J. Biosci. Bioeng.* 91, 251–255.
- Mellado, E., Aufauvre-Brown, A., Specht, C.A., Robbins, P.W., Holden, D.W., 1995. A multigene family related to chitin synthase genes of yeast in the opportunistic pathogen *Aspergillus fumigatus*. *Mol. Gen. Genet.* 246, 353–359.
- Merz, R.A., Horsch, M., Nyhlen, L.E., Rast, D.M., 1999. Biochemistry of chitin synthase. In: Jollés, P., Muzarelli, R.A.A. (Eds.), *Chitin and Chitinases*. Birkhäuser-Verlag, Basel, Switzerland, pp. 9–37.
- Nagahashi, S., Sudoh, M., Ono, N., Sawada, R., Yamaguchi, E., Uchida, Y., Mio, T., Takagi, M., Arisawa, M., Yamada-Okabe, H., 1995. Characterization of chitin synthase of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 270, 13961–13967.
- Nielsen, H., Engelbrecht, J., Brunak, S., Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineer.* 10, 1–6.
- Nielsen, H., Krogh, A., 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. In: *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6)*. AAAI Press, Menlo Park, CA, pp. 122–130.
- Nino-Vega, G.A., Buurman, E.T., Gooday, G.W., San-Blas, G., Gow, N.A., 1998. Molecular cloning and sequencing of a chitin synthase gene (CHS2) of *Paracoccidioides brasiliensis*. *Yeast* 14, 181–187.
- Palli, S.R., Retnakaran, A., 1999. Molecular and biochemical aspects of chitin synthesis inhibition. In: Jollés, P., Muzarelli, R.A.A. (Eds.), *Chitin and Chitosan*. Birkhäuser Verlag, Basel, Switzerland, pp. 85–98.

- Richmond, T. 2000. Higher plant cellulose synthases. *Genome Biol.* 1, REVIEWS3001–3005.
- Rini, J.M., Sharon, N., 2000. Carbohydrates and glycoconjugates: glycosyltransferases, sugar nucleotide transporters and bacterial surface lectins—at the cutting edge of glycobiology. *Curr. Opin. Struct. Biol.* 10, 507–509.
- Sonnhammer, E.L.L., von Heijne, G., Krogh, A., 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. In: Glasgow, J., Littlejohn, T., Major, F., Lathrop, R., Sankoff, D., Sensen, C. (Eds.), *Proc. Sixth Int. Conf. on Intelligent Systems for Molecular Biology*. AAAI Press, p, Menlo Park, CA, pp. 175–182.
- Specht, C.A., Liu, Y., Robbins, P.W., Bulawa, C.E., Iartchouk, N., Winter, K.R., Riggle, P.J., Rhodes, J.C., Dodge, C.L., Culp, D.W., Borgia, P.T., 1996. The *chsD* and *chsE* genes of *Aspergillus nidulans* and their roles in chitin synthesis. *Fungal Genet. Biol.* 20, 153–167.
- Spindler, K.D., Spindler-Barth, M., 1999. Inhibitors of chitinases. In: Jollés, P., Muzzarelli, R.A.A. (Eds.), *Chitin and Chitosan*. Birkhäuser Verlag, Basel, Switzerland, pp. 201–209.
- Tellam, R.L., Vuocolo, T., Johnson, S.E., Jarmey, J., Pearson, R.D., 2000. Insect chitin synthase. cDNA sequence, gene organization and expression. *Eur. J. Biochem.* 267, 6025–6042.
- The *C. Elegans* Sequencing Consortium, 1998. *Science* 282, 2012–2018.
- Unligil, U.M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., Rini, J.M., 2000. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J.* 19, 5269–5280.
- Wiggins, C.A.R., Munro, S., 1998. Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc. Natl. Acad. Sci. U S A* 95, 7945–7950.
- Yabe, T., Yamada-Okabe, T., Nakajima, T., Sudoh, M., Arisawa, M., Yamada-Okabe, H., 1998. Mutational analysis of chitin synthase 2 of *Saccharomyces cerevisiae*. Identification of additional amino acid residues involved in its catalytic activity. *Eur. J. Biochem.* 258, 941–947.
- Yanai, K., Kojima, N., Takaya, N., Horiuchi, H., Ohta, A., Takagi, M., 1994. Isolation and characterization of two chitin synthase genes from *Aspergillus nidulans*. *Biosci. Biotechnol. Biochem.* 58, 1828–1835.
- Zhang, Z.G., Hall, A., Perfect, E., Gurr, S.J., 2000. Differential expression of two chitin synthase genes of *Blumeria graminis*. *Mol. Plant Pathol.* 1, 125–138.